⌐S9C
10/09 **OIPE**

RAW SEQUENCE LISTING   DATE:
PATENT APPLICATION: US/09/851,410   TIME:

Input Set : **N:\Crf3\RULE60\09851410.txt**
Output Set: **N:\CRF3\06152001\I851410.raw**

ENTERED

SEQUENCE LISTING

```
 3  (1) GENERAL INFORMATION:
 5      (i) APPLICANT: Reyes, Gregory R
 6                    Yarbough, Patrice O
 7                    Bradley, Daniel W
 8                    Krawczynski, Krzysztof Z
 9                    Tam, Albert
10                    Fry, Kirk E
11     (ii) TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
13                              Non-A/Non-B Hepatitis Viral Agent
15    (iii) NUMBER OF SEQUENCES: 20
17     (iv) CORRESPONDENCE ADDRESS:
18          (A) ADDRESSEE: Dehlinger & Associates
19          (B) STREET: 350 Cambridge Avenue, Suite 250
20          (C) CITY: Palo Alto
21          (D) STATE: CA
22          (E) COUNTRY: USA
23          (F) ZIP: 94306
25      (v) COMPUTER READABLE FORM:
26          (A) MEDIUM TYPE: Floppy disk
27          (B) COMPUTER: IBM PC compatible
28          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
29          (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
31     (vi) CURRENT APPLICATION DATA:
```
C--> 32   **(A) APPLICATION NUMBER: US/09/851,410**
C--> 33   **(B) FILING DATE: 07-May-2001**
```
39    (vii) PRIOR APPLICATION DATA:
36          (A) APPLICATION NUMBER: 09/128,275
37          (B) FILING DATE:
40          (A) APPLICATION NUMBER: US 07/681,078
41          (B) FILING DATE: 05-APR-1991
44          (A) APPLICATION NUMBER: US
45          (B) FILING DATE: 06-APR-1990
48          (A) APPLICATION NUMBER: US       421,421
49          (B) FILING DATE: 13-OCT-1989
52          (A) APPLICATION NUMBER: US       367,486
          (B) FILING DATE: 16-JUN-1989
          (A) APPLICATION NUMBER: US       
57          (B) FILING DATE: 11-APR-1989
          (A) APPLICATION NUMBER: US       
61          (B) FILING DATE: 17-JUN-1988
65   (viii) ATTORNEY AGENT INFORMATION:
64          (A) NAME: Dehlinger, Peter B.
65          (B) REGISTRATION NUMBER: 41,
66          (C) REFERENCE DOCKET NUMBER: 
68     (ix) TELECOMMUNICATION INFORMATION:
69          (A) TELEPHONE: 
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/851,410

DATE: ......
TIME: ......

Input Set : N:\Crf3\RULE60\09851410.txt
Output Set: N:\CRF3\06152001\I851410.raw

```
 70          (B) TELEFAX: .651  324-.301
 72 (2) INFORMATION FOR SEQ ID NO: 1:
 74     (i) SEQUENCE CHARACTERISTICS:
 75          (A) LENGTH: 1295 base pairs
 76          (B) TYPE: nucleic acid
 77          (C) STRANDEDNESS: double
 78          (D) TOPOLOGY: linear
```

W--> 80   (ii) MOLECULE TYPE: DNA

```
 82    (iii) HYPOTHETICAL: NO
 84     (iv) ANTI-SENSE: NO
 86     (vi) ORIGINAL SOURCE:
 87          (C) INDIVIDUAL ISOLATE: 1.36 kb EcoRI insert of ET1.1,
 88                                   forward sequence
 90     (ix) FEATURE:
 91          (A) NAME/KEY: CDS
 92          (B) LOCATION: 1..1293
 94     (ix) FEATURE:
 95          (A) NAME/KEY: CDS
 96          (B) LOCATION: 2..1294
 97     (ix) FEATURE:
 98          (A) NAME/KEY: CDS
100          (B) LOCATION: 3..1295
103     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
105 AGACCTGTCC CTGTTGCAGC TGTTCTACCA CCCTGCCCCG AGCTCGAACA GGGCCTTCTC      60
107 TACCTGCCCC AGGAGCTCAC CACCTGTGAT AGTGTCGTAA CATTTGAATT AACAGACATT     120
109 GTGCACTGCC GCATGGCCGC CCCGAGCCAG CGCAAGGCCG TGCTGTCCAC ACTCGTGGGC     180
111 CGCTACGGCG GTGGCACAAA GCTCTACAAT GCTTCCCACT CTGATGTTCG CGACTCTCTC     240
113 GCCCGTTTTA TCCCGGCCAT TGGCCCCGTA CAGGTTACAA CTTGTGAATT GTACGAGCTA     300
115 GTGGAGGCCA TGGTCGAGAA GGGCCAGGAT GGCTCCGCCG TCCTTGAGCT TGATCTTTGC     360
117 AACCGTGACG TGTCCAGCAT CACCTTCTTC CAGAAAGATT GTAACAAGTT CACCACAGGT     420
119 GAGACCATTG CCCATGGTAA AGTGGGCCAG GGCATCTCGG CCTGGAGCAA GACCTTCTGC     480
121 GCCCTCTTTG GCCCTTGGTT CCGGCGTATT GAGAAGGCTA TTCTGGCCCT GCTCCCTCAG     540
123 GGTGTGTTTT ACGGTGATGC CTTTGATGAC ACCGTCTTCT CCGCGGCTGT GGCCGCAGCA     600
125 AAGGCATCCA TGGTGTTTGA GAATGACTTT TCTGAGTTTG ACTCCACCCA GAATAACTTT     660
127 TCTCTGGGTC TAGAGTGTGC TATTATGGAG GAGTGTGGGA TGCCGCAGTG GCTCATCCGC     720
129 CTGTATCACC TTATAAGGCC TGCGTGGATG TTGCAGGCCC CGAAGGAGTC TGTGCGAGGG     780
131 TTTTGGAAGA AACACTCCGG TGAGCCTGGC ACTCTTCTAT GGAATACTGT CTGGAATATG     840
133 GCCGTTATTA CCCATTGTTA TGACTTCCGC GATTTTCAGG TGGCTGCCTT TAAAGGTGAT     900
135 GATTCGATAG TGCTTTGCAG TGAGTATGGT CAGAGTCCAG GAGCTGCTGT CCTGATCGCC     960
137 GGCTGTGGCT TGAAGTTGAA GGTAGATTTC CGGCCGATCG GTTTGTATGC AGGTGTTGTG    1020
139 GTGGCCCCCG GCCTTGGCGC GCTTCCTGAT GTTGTGAGCT TCGGCGGCCG GCTTACCGAG    1080
141 AAGAATTGGG GCCCTGCCC TGAGCGGGCG GAGCAGCTCA GCCTCGCTGT TAGTGATTTC    1140
143 CCCGGCAAGC TCACGAATGT AGCTCAGATG TGTGTGATG TTGTTCCCG TGTTTATGCG     1200
145 GTTTCCCCTG GACTGGTTCA TAACCTGATT GGCATGGTAT AGGTGTTGC TTATGGCAAG    1260
147 GCACATTTCA CTGAGTCAGT AAAACGAGTG CTCCA                               1295
150 (2) INFORMATION FOR SEQ ID NO: 2:
152     (i) SEQUENCE CHARACTERISTICS:
153          (A) LENGTH: 431 amino acids
154          (B) TYPE: amino acid
```

```
155              TOPOLOGY: linear
156      (ii) MOLECULE TYPE: protein
159      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
161  Arg Pro Val Pro Val Ala Ala Val Leu ... Tyr Pro ... Leu ...
163    1           5                             15
164  Gln Gly Leu Leu Tyr Leu Pro Gln ... Thr Thr Tyr Asp ... Val
166          20              25
167  Val Thr Phe Glu Leu Thr Asp Ile Val His Lys Arg Met Ala Ala ...
169          35              40              45
170  Ser Gln Arg Lys Ala Val Leu Ser Thr Leu Val Gly Arg Tyr Gly Gly
172    50              55              60
173  Arg Thr Lys Leu Tyr Asn Ala Ser His ... Asp Val Arg Asp Ser Leu
174   65              70              75              80
176  Ala Arg Phe Ile Pro Ala Ile Gly Pro Val ... Val Thr Thr Cys Glu
177              85              90              95
179  Leu Tyr Glu Leu Val Glu Ala Met Val Glu Lys Gly Gln Asp Gly Ser
180             100             105             110
181  Ala Val Leu Glu Leu Asp Leu Cys Asn Arg Asp Val Ser Arg Ile Thr
182             115             120             125
185  Phe Phe Gln Lys Asp Cys Asn Lys Phe Thr Thr Gly Glu Thr Ile Ala
186             130             135             140
188  His Gly Lys Val Gly Gln Gly Ile Ser Ala Trp Ser Lys Thr Phe Cys
189  145             150             155             160
191  Ala Leu Phe Gly Pro Trp Phe Arg Ala Ile Glu Lys Ala Ile Leu Ala
192             165             170             175
194  Leu Leu Pro Gln Gly Val Phe Tyr Gly Asp Ala Phe Asp Asp Thr Val
195             180             185             190
197  Phe Ser Ala Ala Val Ala Ala Ala Lys Ala Ser Met Val Phe Glu Asn
198             195             200             205
200  Asp Phe Ser Glu Phe Asp Ser Thr Gln Asn Asn Phe Ser Leu Gly Leu
201         210             215             220
203  Glu Cys Ala Ile Met Glu Glu Cys Gly Met Pro Gln Trp Leu Ile Arg
204  225             230             235             240
206  Leu Tyr His Leu Ile Arg Ser Ala Trp Ile Leu Gln Ala Pro Lys Glu
207             245             250             255
209  Ser Leu Arg Gly Phe Trp Lys Lys His Ser Gly Glu Pro Gly Thr Leu
210             260             265             270
211  Leu Trp Asn Thr Val Trp Asn Met Ala Val Ile Thr His Cys Tyr Asp
212             275             280             285
214  Phe Arg Asp Phe Gln Val Ala Ala Phe Lys Gly Asp Asp Ser Ile Val
216             290             295             300
217  Leu Cys Ser Glu Tyr Arg Gln Ser Pro Gly Ala Ala Val Leu Ile Ala
218  305             310             315             320
221  Gly Cys Gly Leu Lys Leu Lys Val Asp Phe Arg Pro Ile Gly Leu Tyr
222             325             330             335
224  Ala Gly Val Val Val Ala Pro Gly Leu Gly Ala Leu Pro Asp Val Val
225             340             345             350
227  Arg Phe Ala Gly Arg Leu Thr Thr Lys Asn Trp Gly Pro ... Gly ...
225             355             360             365
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/851,410

DATE: ...
TIME: ...

Input Set : N:\Crf3\RULE60\09851410.txt
Output Set: N:\CRF3\06152001\I851410.raw

```
290 Arg Ala Glu ... ... Arg Leu Ala Val ... Asp Ile Leu Arg Lys Leu
291     370              375                 ...
293 Thr Asn Val Ala Gln Met Lys Val Asp Val ... Ser Arg Val Tyr Gly
...  385              390               ...              ...
... Val Ser Pro Gly Leu Val His Asn Leu ... Gly Met Leu Gln Ala Val
...                   405              ...                    415
... Ala Asp Gly Lys Ala His Phe Thr Glu ... Val Lys Pro Val Leu
...                   420              425                    430
... (2) INFORMATION FOR SEQ ID NO: 3:
...      (i) SEQUENCE CHARACTERISTICS:
...          A) LENGTH: 18 base pairs
...          B) TYPE: nucleic acid
...          C) STRANDEDNESS: single
...          D) TOPOLOGY: linear
```

W--> 251   (ii) MOLECULE TYPE: DNA
```
...    (iii) HYPOTHETICAL: NO
...    (iv) ANTI-SENSE: NO
...    (vi) ORIGINAL SOURCE:
...          C) INDIVIDUAL ISOLATE: linker - top (5') sequence
...    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
... GGAATTCGCG GCGGCTCG                                              18
... (2) INFORMATION FOR SEQ ID NO: 4:
...      (i) SEQUENCE CHARACTERISTICS:
...          A) LENGTH: 20 base pairs
...          B) TYPE: nucleic acid
...          C) STRANDEDNESS: single
...          D) TOPOLOGY: linear
```

W--> 273   (ii) MOLECULE TYPE: DNA
```
...    (iii) HYPOTHETICAL: NO
...    (iv) ANTI-SENSE: NO
...    (vi) ORIGINAL SOURCE:
...          C) INDIVIDUAL ISOLATE: linker - bottom (3') sequence
...    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
... CGAGCGGCCG CGAATTCCTT                                            20
... (2) INFORMATION FOR SEQ ID NO: 5:
...      (i) SEQUENCE CHARACTERISTICS:
...          A) LENGTH: 1235 base pairs
...          B) TYPE: nucleic acid
...          C) STRANDEDNESS: double
...          D) TOPOLOGY: linear
```

W--> 294   (ii) MOLECULE TYPE: DNA
```
...    (iii) HYPOTHETICAL: NO
...    (iv) ANTI-SENSE: NO
...    (vi) ORIGINAL SOURCE:
...          C) INDIVIDUAL ISOLATE: 1.38 kb EcoRI insert of ET1.1,
...                                 reverse sequence
...    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
... TCGAGCACTG CTTTTACTCA CTCACTGAAA TCT... ...ATCA...AAT A...TCTAACT    60
... ATGCCAATCA ...TTATCAAT GACTTCAGCT CAAC...TAT AAACA...CCA ACAAACATCC   120
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/851,410

DATE: ~~~~~~
TIME: ~~~~~~

Input Set : N:\Crf3\RULE60\09851410.txt
Output Set: N:\CRF3\06152001\I851410.raw

```
311 ACACACATCT CAGGCTACATT CGTCAGCTTG CGGATGAAAT CACTAACAGC GAGGCGGAGC      180
312 TGTTTGGTCG CGTCAGGGCC AGGGCGGCAA ~~~~~~~~~~ TAAGCCGGCC GCGGAAGGTG      140
31~ ACAACATTAG GCAGCGCGCC AAGGCCGGGG GCCATCATAA CACCTGCATA CAAACCGATC      300
31~ GGGCGGAAAT CTACCTTCAA CTTCAAGCCA CAGCGGCGGA TCAGGACAGC AGTTCCTGGA      360
314 CTCCGAGGAT ACTCACTGCA AAGCACTATC GAATCATCAC CTTTAAAGGC AGCCACGTGA      420
321 AAATTGCGGA ACTCATAACA GTGGGTAATA ACGGCGATAT TCCAGACAGT ATTCCATAGA      480
313 AGAGTGCGGG GCTCAGCGGA GTGTTCTTC CAAAATCCTG GCAGAGACTC CTTCGGGGCC      540
31~ TGGAAGATCC ATGCAGACCT TATAAGGTGA TACAGGCGGA TGAGCCACTG CGGCATCCCA      600
31~ CATTTCTTCA AAATAGCACA CTGCACACCC AGAGAAAAGT TATTCTGGGT GGAGTCAAAC      660
31~ TCAGAAAAGT TATTCTCAAA CAGCATGGAT GCGTTTGCTG CGGCCACAGC CGCCGAGAAG      720
3~1 ACGGTGTTAT GAAAGGCATC ACGGTAAAAC ACACGGTGAG GGAGCAGGGC CAGAATAGCC      780
3~3 TTCTCAATAG GGCGGAACGA AGGGCTAAAG AGGGCGCAGA AGGTCTTGCT CCAGGCCGAG      840
3~~ ATGCGCTGGG CCACTTTACC ATGGGTAATG GTCTCACCTG TGGTGAACTT GTTACAATCT      900
3~~ TTCTGCAAGA AGGTGATGCT GGACAGGTCA CGGTTGCAAA GATCAAGCTC AAGGACGGCG      960
3~~ GAGCCATTCT GGCGCTTCTC GACCATGGCG TCCACTAGCT CGTACAATTC ACAAGTTGTA     1020
341 ACCTGTAGGG GGCGAATGGC CGGGATAAAA CGGGCGAGAG AGTCGCGAAC ATCAGAGTGG     1080
343 GAAGCATTGT AGAGCTTTGT GCGACCGCCG TAGCGGCCCA CGAGTGTGGA CAGCACGGCC     1140
34~ TTGCGCTGGC TCGGGGCGGG CATGCGGCAG TGCACAATGT CTGTTAATTC AAATGTTACG     1200
347 ACACTATCAG AGGTGGTGAG TTCCTGGGGC AGGTAGAGAA GGCCCTGTTC GAGCTCGGGG     1260
                                                                          1295
3~9 CAGGGTCGTA GAACAGCTGG AACAGGCACA GGTCT
```

3~2 (2) INFORMATION FOR SEQ ID NO: 6:
3~4   (i) SEQUENCE CHARACTERISTICS:
3~5   (A) LENGTH: 1195 base pairs
3~6   (B) TYPE: nucleic acid
W--> 367   (C) STRANDEDNESS: HEV - Burma strain
3~8   (D) TOPOLOGY: linear
W--> 360   (ii) MOLECULE TYPE: DNA
3~2   (iii) HYPOTHETICAL: N
3~4   (iv) ANTI-SENSE: NO
3~6   (vi) ORIGINAL SOURCE:
3~9   (ix) FEATURE:
3~0   (A) NAME/KEY: CDS
3~1   (B) LOCATION: 28..5106
3~3   (ix) FEATURE:
3~4   (A) NAME/KEY: CDS
3~~   (B) LOCATION: 5147..7126
3~~   (ix) FEATURE:
3~~   (A) NAME/KEY: CDS
3~4   (B) LOCATION: 5106..5474
3~2   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
3~4 AGGCAGCGCA CATATGTGGT TCATGGCATG GAGGCCCATC AGTTTATTAA GGCTCCTGGC       60
3~6 ATCACTACTG CTATTGAGCA GGCTGTTCTA GCAGCGGCCA ACTCTGCCCT GGCGAATGCT      120
3~8 GTGGTAGTTA GGCCTTTTGT TTCTCACCAG CAGATTGAGA TCCTCATTAA CCTAATGCAA      180
3~0 CCTGGTGAGC TTGTTTTCGG CGGCGAGGTT TTCTGGAATC ATCCCATCCA GCCGTGTCATC      240
3~2 CATAACTAGC TGGAGCTTTA CTGCCGCGCC CGGTCCGGCC GCTGTCTTGA AATTGGCGCC      300
3~4 CATCCGGGT CAATAAATGA TAATCCTAAT GTGGTGCACC GCTGCTTCCT CCGCCCTGTT      360
3~6 GGGGTGATG TTCAGCGCTG GTATACTGCT GAGCGCGATT ACTAGTCGA CGGGTTTTCT      420
398 TGTTCTGCGC TGCGCGGCTT TGCCGCTGCT GAGGCGGATT ACTAGTCGA CGGGTTTTCT      480
4~0 GGCTGTAACT TTTGCGGGCA GACTGGGATC CATTTTTATT GGTTGATGA TATGTTAGTA      540
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/851,410

DATE: . : . . :
TIME: 1: : : . 6

Input Set : **N:\Crf3\RULE60\09851410.txt**
Output Set: **N:\CRF3\06152001\I851410.raw**

```
L:32 M:220 C: Keyword misspelled or invalid format, [ A] APPLICATION NUMBER:]
L:33 M:220 C: Keyword misspelled or invalid format, [ B  FILING DATE:]
L:80 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:251 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3
L:273 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4
L:294 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5
L:367 M:21  C: Keyword misspelled or invalid format, [ M  STRANDEDNESS:]
L:367 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=6
L:361 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:1142 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1142 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=10
L:1135 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10
L:1465 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1465 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=11
L:1548 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11
L:1477 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1477 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=12
L:1470 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12
L:1651 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1651 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=17
L:1680 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1680 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=18
L:1708 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1708 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=19
L:1736 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1736 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=20
```